September 2015

Benjamin Erhardt

Choicing
A choice-based method to create a ranking

Preface

The human brain is not capable of ranking, but only of taking choice decisions. For this reason, it is a challenging task to create a ranking of an unordered set of elements. Our brain naturally breaks the task down to choice decisions, e.g. by selecting the most preferred element, assign this element to the highest available rank, then repeat, until the ranking is complete.
It is a known biasing problem in the full profile approach of Conjoint Analysis: rankings, which have been created by persons without any help, often do not represent the actual preferences of the person, especially for the „middle“ ranks (the ranks which are not perceived as best or worst couple of ranks).
The goal of this paper is to present a method „Choicing“ which provides a fast and simple “choice-based” approach to create rankings of elements using a sequence of simple choice decisions. These “choice-based” rankings should perfectly represent the persons’ preferences for the elements.

A ranking as the result of its elements‘ preferences

We have a number of elements E. Each element e_i with i = 1, 2, …, N has a rank r_i. The ranks R are ordinally scaled. Usually, the rank r = 1 represents the most preferred resp. best rank, the following ranks are incremented by 1, so the 2nd best rank has r = 2, the 3rd best rank r = 3 etc. Once the ranking is finished, each rank r is unique (so no „tied“ ranks anymore).
This implies that the ranking R for elements E is a representation of all preferences P, with p_{ij} being the preference between e_i and e_j.

There are 3 possible manifestations of p_{ij}:

$p_{ij} = 0$ | Indifference (no preference known yet)

$p_{ij} = 1$ | Preference of e_i over e_j (or: $e_i > e_j$)

$p_{ij} = -1$ | Preference of e_j over e_i (or: $e_i < e_j$)

If no ranks have been assigned yet, all p_{ij} are 0.

If all ranks have been assigned, then all p_{ij} are 1 or -1 (with $i \neq j$).

Let us visualize these relations using a simple example:
We have 5 elements e_1 = A, e_2 = B, e_3 = C, e_4 = D and e_5 = E.

So the „empty“ ranking and its preferences look like this:

	Ranks	Elements	Preferences					
i	**r**	**e**	**p**	1	2	3	4	5
1		A	1	0	0	0	0	0
2		B	2	0	0	0	0	0
3		C	3	0	0	0	0	0
4		D	4	0	0	0	0	0
5		E	5	0	0	0	0	0

When all preferences are known and thus all ranks are assigned, the tables look like this (here, we assume preferences of A > B > C > D > E):

	Ranks	Elements	Preferences					
i	**r**	**e**	**p**	1	2	3	4	5
1	1	A	1	0	1	1	1	1
2	2	B	2	-1	0	1	1	1
3	3	C	3	-1	-1	0	1	1
4	4	D	4	-1	-1	-1	0	1
5	5	E	5	-1	-1	-1	-1	0

The relation between an element’s preferences and its rank

To understand the strict relation of each rank r_i to the preferences p_{ij}, we need to introduce the score s_i. The score s_i of element e_i is determined as follows:

$s_i = \sum_j p_{ij}$

Now, we assign rank $r_i = 1$ to the maximum s_i, $r_i = 2$ to the 2nd highest s_i etc. If $s_i = s_j$, we speak of tied ranks, and thus $r_i = r_j$. The next rank is not incremented by 1 in such cases, but by the number of elements which have a tied rank.

The following illustration shows the relation between r_i and s_i for the finished ranking of our example:

	Ranks	Elements	Preferences						Score
i	**r**	**e**	**p**	1	2	3	4	5	**s**
1	**1**	A	1	0	1	1	1	1	**4**
2	**2**	B	2	-1	0	1	1	1	**2**
3	**3**	C	3	-1	-1	0	1	1	**0**
4	**4**	D	4	-1	-1	-1	0	1	**-2**
5	**5**	E	5	-1	-1	-1	-1	0	**-4**

Determination of final ranks by open preferences

The above example represents a finished ranking because all ranks are final. Formally, all ranks are final because all p_{ij} are 1 or -1 (with i ≠

j). So in other words, a rank r_i is final when the number of indifferent preferences resp. „open" choices $o_i = 0$.

Adding the open choices to our visualization of the finished ranking provides:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
1	1	A	1	0	1	1	1	1	4	**0**
2	2	B	2	-1	0	1	1	1	2	**0**
3	3	C	3	-1	-1	0	1	1	0	**0**
4	4	D	4	-1	-1	-1	0	1	-2	**0**
5	5	E	5	-1	-1	-1	-1	0	-4	**0**

Let us now assume that the preference of D > E is not yet known;
so $p_{45} = 0$ instead of $p_{45} = 1$, and $p_{54} = 0$ instead of $p_{54} = -1$.
Now, our table looks like this:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
1	1	A	1	0	1	1	1	1	4	0
2	2	B	2	-1	0	1	1	1	2	0
3	3	C	3	-1	-1	0	1	1	0	0
4	**4**	D	4	-1	-1	-1	0	**0**	**-3**	**1**
5	**4**	E	5	-1	-1	-1	**0**	0	**-3**	**1**

This has a couple of implications:

1. The ranks $r_4 = r_5 = 4$ are tied, as the scores $s_4 = s_5 = -3$ are the same.
2. The ranks r_4 and r_5 are not final, as $o_4 > 0$ and $o_5 > 0$.
3. The ranking is not finished, as not all $o_i = 0$.

Choice decisions to update the ranking

Again, we start with an empty ranking:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
1	1	A	1	0	0	0	0	0	0	4
2	1	B	2	0	0	0	0	0	0	4
3	1	C	3	0	0	0	0	0	0	4
4	1	D	4	0	0	0	0	0	0	4
5	1	E	5	0	0	0	0	0	0	4

Naturally, in an empty ranking all $o_i = 4$, as in a set of 5 elements, each element has 4 preference relations with other elements; so, before taking decisions each element has 4 open choices.

Now, we present a set of elements, of which the preferred element should be chosen. We assume the set of elements would be A, B, C; and element e_1 = A would be chosen as the preferred element (indicated by *):

- A* (e_1)
- B (e_2)
- C (e_3)

This implies 4 preferences:

- p_{12} = 1
- p_{13} = 1
- p_{21} = -1
- p_{31} = -1

We write these preferences in our table view to receive:

	Ranks	Elements		Preferences						Score	Open
i	r	e		p	1	2	3	4	5	s	o
1	**1**	A		1	0	**1**	**1**	0	0	**2**	**2**
2	**4**	B		2	**-1**	0	0	0	0	**-1**	**3**
3	**4**	C		3	**-1**	0	0	0	0	**-1**	**3**
4	**2**	D		4	0	0	0	0	0	**0**	**4**
5	**2**	E		5	0	0	0	0	0	**0**	**4**

Summary:

1. The 4 preferences are in the preferences matrix.
2. The scores imply the ranks (with tied ranks for same scores).
3. None of the ranks is final yet, as all $o_i > 0$.

Limit to sets of 2 or 3 elements

As already mentioned, our brain needs more time to choose if the number of elements increases. A real instant decision can be taken in cases 1 out of 2, and 1 out of 3. If we have to choose from 4 or more elements, we start to iterate smaller choice decisions. As this is tiring and inefficient, we limit the number of elements in the choice sets to 3 elements. As there are cases, where the next most efficient choice decision requires only a choice 1 out of 2 (e.g. if only the choice decision between exactly 2 elements is “open”), a choice set is also allowed to have 2 elements only.

Priority 1: highest score, priority 2: lowest open

In Choicing, it is about reaching the next most efficient ranking as fast as possible. This refers to the fact, that each ranking on the way to the actual final ranking is relevant to the person who creates the ranking. So, if the person cancels before reaching the final ranking, he or she wishes to have the most preferred elements first.

To respect this requirement, each next set needs to represent the 3 (or 2) elements e_i with the highest scores s_i which are not final resp. have open choices ($o_i > 0$).

If there are more than 3 possible elements, we try to focus on reaching the final state as fast as possible. Thus, we choose the 3 (or 2) elements with the lowest (number of) open (choices) o_i.

If this still makes more than 3 elements possible, we randomly choose from the elements with tied lowest scores.

In our example, the next set is found by selecting the 3 highest score, which returns exactly 3 results:

	Ranks	Elements	Preferences						Score	Open
i	r	e	p	1	2	3	4	5	s	o
1	**1**	**A**	**1**	**0**	**1**	**1**	**0**	**0**	**2**	**2**
2	4	B	2	-1	0	0	0	0	-1	3
3	4	C	3	-1	0	0	0	0	-1	3
4	**2**	**D**	**4**	**0**	**0**	**0**	**0**	**0**	**0**	**4**
5	**2**	**E**	**5**	**0**	**0**	**0**	**0**	**0**	**0**	**4**

So, the next choice set:

- A (e_1)
- D (e_4)
- E (e_5)

Strict Preferences

In order to make sure to always reach a unique rank r_i for each element e_i, we apply strict preferences.

The idea of strict preferences in a nutshell (with > for „is preferred over“):

If A > B and B > C, then A > C

Let us assume that next choice would be D:

- A (e_1)
- D* (e_4)
- E (e_5)

This implies 4 (direct) preferences:

- p_{41} = 1 (D > A)
- p_{45} = 1 (D > E)
- p_{14} = -1 (A < D)
- p_{54} = -1 (E < A)

Applying strict preference for D > A, this implies another 4 preferences:

- p_{42} = 1 (D > B, as D > A and A > B)
- p_{43} = 1 (D > C, as D > A and A > C)
- p_{24} = -1 (B < D, as D > A and A > B)
- p_{34} = -1 (C < D, as D > A and A > C)

So our updated ranking looks like this:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
1	2	A	1	0	1	1	**-1**	0	1	1
2	4	B	2	-1	0	0	**-1**	0	-2	2
3	4	C	3	-1	0	0	**-1**	0	-2	2
4	1	D	4	**1**	**1**	**1**	0	**1**	4	0
5	3	E	5	0	0	0	**-1**	0	-1	3

Summary:

1. The rank $r_4 = 1$ is final, as $o_4 = 0$ (also as $s_4 = 4$ which is the maximum possible score for 5 elements)
2. For the next set, all remaining 4 elements are possible, as we have a tie for the 3rd element; both B and C have s = -2 and o = 2, so in this situation, we choose randomly. Assuming the random choice would provide B, our next choice set would be (with choice of E):
   - A (e_1)
   - B (e_2)
   - E* (e_5)

So again, strict preferences implies 2 additional preferences from E > A and A > C, which adds E > C:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
1	**3**	A	1	0	1	1	-1	**-1**	**0**	**0**
2	4	B	2	-1	0	0	-1	**-1**	-3	1
3	4	C	3	-1	0	0	-1	**-1**	-3	1
4	**1**	D	4	1	1	1	0	1	**4**	**0**
5	**2**	E	5	**1**	**1**	**1**	-1	0	**2**	**0**

Now, the ranking is nearly finished, as the ranks $r_4 = 1$, $r_5 = 2$ and $r_1 = 3$ are final, indicated by $o_4 = o_5 = o_1 = 0$.

The final „choice-based“ ranking

In the example, the last choice decision is a set with only 2 elements (we assume choice of B):

- B* (e_2)
- C (e_3)

This creates the following final ranking. For the first time, we present the ranking sorted for rank which would be the way to always present the ranking in a user interface for people applying „Choicing“ to create rankings:

	Ranks	Elements	Preferences						Score	Open
i	**r**	**e**	**p**	1	2	3	4	5	**s**	**o**
4	**1**	**D**	4	1	1	1	0	1	4	0
5	**2**	**E**	5	1	1	1	-1	0	2	0
1	**3**	**A**	1	0	1	1	-1	-1	0	0
2	**4**	**B**	2	-1	0	1	-1	-1	-2	0
3	**5**	**C**	3	-1	-1	0	-1	-1	-4	0